Applied Human Factors and Ergonomics (AHFE2025), Vol. 199, 2025, 1214–1220
https://doi.org/10.54941/ahfe1006932

Measuring Group Cohesion as a Factor in Collaborative Decision Making

Tomek Strzalkowski

Tetherless World Constellation Department of Cognitive Science Rensselaer Polytechnic Institute Troy, NY, USA

ABSTRACT

We have developed a collection of AI tools that can estimate the degree of group cohesion by measuring the quality of interaction between group members. To compute this assessment, we identify a series of sociolinguistic behaviors, both individual and collective, that obtain within the group as reflected in the members' utterances. Individual behaviors impacting group cohesion include agenda control, involvement, agreement and disagreement; relevant collective behaviors include the balance of agreement-disagreement, sociability, and task focus. Estimating distribution and degree of these behaviors are required to assess the character of members' relations and the consistency of group's objective. In this paper we focus on collective behaviors, and how they can be automatically computed from group interactions. Implementation of individual behaviors can be found in (Broadwell et al., 2013). In this paper we discuss a design of an AI system that can determine cohesiveness of a group engaged in task-oriented discussion.

Keywords: Social dynamics in conversation, Group cohesion, Sociolinguistic behaviors

INTRODUCTION

We define group cohesion as a general consistency in group objectives or values over time that encompasses an overall persistence in the cooperative nature of the members' interactions. This characterization applies primarily to task-oriented teams, although it may also be used more broadly to purpose-driven groups. One measure of group cohesion is a degree of task focus among the group members. It is a collective behavior that can be measured by the degree to which the discussion is focused on *a shared objective*, as well as by the efficiency with which the group works towards this objective. The efficiency of this progress is evidenced by the degree to which the discourse stays on topic with few off-topic digressions. Another measure of group cohesion is *persistence of roles* (Bales, 1950; 2001), which tracks whether certain key social functions in the group, such as leadership, persist throughout the discourse, though not necessarily filled continuously by the same individuals. A cohesive group is also characterized by a high degree of *sociability*. This includes adherence to general conversational principles, which are in turn reflected by certain sequences of dialogue acts, i.e., question-answer, offer-response, as well as sequences of expressions classified

as conversational norms, including greetings, thanks and apologies. Groups with higher values of the sociability behavior are considered more cohesive.

In Broadwell et al. (2013) we introduced a methodology for automatically identifying and quantifying certain types of sociolinguistic behaviors in group interactions based solely on the recording of conversations among the group members. Our method applies equally to transcripts of spoken conversations and to text-based chat. Although we primarily focus on synchronous conversations, the method applies also to asynchronous interactions such as these occurring in social media and via email.

The method can be thought of as a hierarchical classifier, where the bottom layers extract basic linguistic indicators from the participants' utterances, such as topics mentioned, vocabulary used, dialogue acts (such as statements, questions, disagreements, etc.), as well as the number of turns per participant and at to whom they are directed. These indicators are subsequently aggregated across longer spans of conversation to reveal topic chains, who initiated them, how long they persist, and who contributes to them. Also revealed are attention and information flow between the participants, including ratios of agreements and disagreements, acceptances and rejections, support and resistance, all of which can range from overt and direct to subtle and indirect.

The aggregated linguistic indicators give raise, in turn, to sociolinguistic behaviors, such as topic control, involvement, network centrality, among others. For example, the rate of topic introduction by a participant, combined with the rate of uptake of these topics by other participants provide a strong basis for estimating the degree of topic control in the conversation. In general, the participants who introduce more topics into the conversation and whose topics are more likely taken up by others have a greater degree of topic control, and consequently more influence in the interaction. This is because topic control is a sociolinguistic behavior that has been shown to correlate with the amount of influence one has in a group. Other sociolinguistic behaviors can be computed in a similar manner, and then combined to establish social roles for individual participants: influence, leadership, power, etc. The overall model is shown in Figure 1. Further details can be found in (Strzalkowski et al., 2012; 2013).

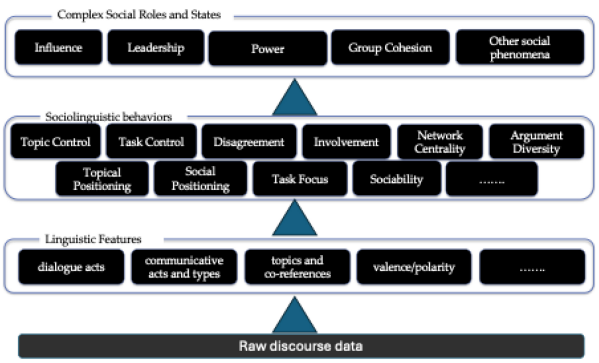

Figure 1: The hierarchical model of sociolinguistic phenomena in group interactions (adapted from Strzalkowski et al., 2012).

COLLECTIVE BEHAVIORS

We focus on computing collective social behaviors, that is, the behaviors that can only be attributed to a group of participants in a conversation, and that require coordinated (though not necessarily pre-planned) activity of all members. A collective behavior can be attributed to an entire group or a subgroup, and it arises as a sum of activities by its members. Group cohesion is a collective phenomenon, which we define as a general consistency in group objectives or values over time and persistence in the cooperative nature of the members' interactions. In other words, a cohesive group could be formed around shared values or a joint task. In this paper we are primarily focusing on task-oriented groups. In particular, we are interested in finding a correlation between (a degree of) group cohesion and the task performance. Thus, the first step is to design a classifier that can decide if a group is cohesive (and, later, to what degree) given a sufficient sample (or samples) of interactions between its members. The groups we investigated in our study ranged from 3 to 8 participants (Shaikh et al., 2010).

ELEMENTS OF GROUP COHESION

Following the design in Broadwell et al. (2013), we identify a set of sociolinguistic behaviors that impact group cohesion (Broadwell et al., 2012). As may be expected these are *collective* behaviors that arise from certain combinations of individual behaviors. The following collective behaviors have been identified as key elements required to operationalize the concept of group cohesion.

Task Focus. One factor of group cohesion is a degree of *task focus* among the group members. Task focus is a collective sociolinguistic behavior that is measured by the degree to which the group discussion is focused on *a shared objective*, as well as by the efficiency with which the group works towards this objective. The presence of a shared objective is evidenced by the largely *single-threaded structure of the most persistent topics in the conversation* (which we call the *meso-topics*), reflecting participants' efforts to keep the discourse progressing towards the objective. The efficiency of this progress is indicated by the degree to which the discourse stays on topic *with few off-topic digressions*. Our hypothesis is that a higher degree of task focus indicates a more cohesive group.

Persistence of Discourse Roles. Another factor of group cohesion is a degree to which the participants engage in certain discourse roles and activities (Bales, 2010). While these roles are expected to persist throughout the discourse, they may not be continuously filled by the same individuals. In this study, we equate discourse roles with high degrees of certain individual behaviors by the group members, e.g., high degree of topic control (= influential); high degree of task control (= leading); high degree of disagreement (= power); etc. The "high degree" of a behavior by a speaker is understood here as significantly greater than the next ranked speaker, e.g., 20% or more. If the next speaker is closer than 20%, then he/she is also assigned this role, and so on. The degree of this collective behavior in the group is the average ratio of speakers with assigned roles across the entire

length of the conversation. Our hypothesis is that a group with a greater PDR value is more cohesive.

Participation. The basic premise of group cohesion is the existence of relationships between the members, i.e., that they do not form an accidental gathering. These relationships are reflected in the amount of interaction among the members, which may be estimated through the combined degree of their individual behaviors such as involvement, agreement, network centrality, etc., that may not rise to level of discourse roles (as defined above) but nonetheless, contribute to the communicative density between the members. A member's degree of participation in the group is determined by the sum of individual behaviors calculated for this member. The group is more cohesive if the mean degree of participation for its members is higher.

Sociability. A cohesive group may be characterized by a high degree of sociability. This includes adherence to general conversational principles, which are reflected by certain sequences of dialogue acts, i.e., question-answer, offer-response, as well as sequences of expressions classified as conversational norms, including greetings, thanks and apologies. For example, when requests for information are (nearly) always answered; directives issued by a speaker are (nearly) always responded to (whether they are followed or not), the degree of sociability between the speakers is high. In addition, high densities of communication links between the participants manifested by either directly addressed utterances or by utterances citing each other topics, indicates greater sociability within the group. When the distribution of these communication links is more balanced between all speakers, the group sociability is greater. Our hypothesis is that groups with a higher value of the sociability measure are more cohesive.

The above design is depicted schematically in Figure 2. Each of the component behaviors is computed on a continuous [0, 1] scale that indicates the degree of each behavior in the group. These scores are then combined (currently using a weighted linear function) to obtain a degree of group cohesion. The model can be further calibrated using supervised machine learning.

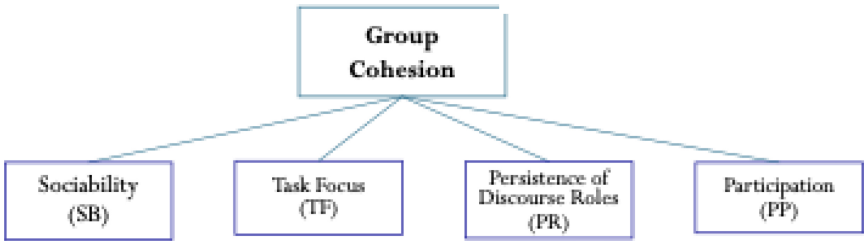

Figure 2: A group cohesion model based on four collective behaviors.

In order to complete the above model, we need to append the linguistic layer that would allow estimating the component collective behaviors directly from the participants' conversations. We explain below how the *sociability* collective behavior is computed from linguistic indicators collected from

the group conversation. Other collective behaviors can be implemented analogously. Sociability is based on the following conversational indicators:

Conversational Norms (observation of): a degree to which the participants adhere to conversational principles such as pairing of questions with answers, offers with responses, etc. It is computed by counting the number of paired vs. unpaired dialogue acts.

Agreement Disagreement (balance of): a degree to which the levels of agreement and disagreement among participants are balanced.

Network Density: the average density of communicative links and topic citations between speakers. We count all direct person-person references and citations of the topics raised by others.

Cite Disparity: an average imbalance of citation links between the speakers.

As may be expected, higher values of the first three measures contribute to the higher degree of sociability, while the high degree of cite disparity would lower it. The four factors are linearly combined with weights calibrated through a machine learning process. Figure 3 shows the corresponding part of the model.

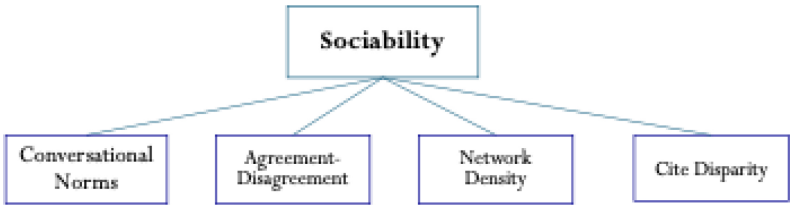

Figure 3: The sociability model based on linguistic indicators in conversation.

Discussion and Further Work

The sociolinguistic model (Fig. 1) has been evaluated for selected aspects, including the majority of individual sociolinguistic behaviors and the roles such as leadership, influence, and power. Details are reported elsewhere (Strzalkowski et al., 2012; 2013; Shaikh et al., 2021, also Katsios et al., 2019). We note here that the model performance has been very good, ranging from 73% to 91% of accuracy for the tested phenomena across multiple communication modes and in two languages (English and Mandarin). Group cohesion was not yet properly evaluated, except for relatively small-scale binary classification (cohesive/not cohesive) which produced 100% accuracy. The main impediment to obtaining reliable test results is the lack of properly annotated ground truth data, and an excessive cost of acquiring such data. Instead, we propose to use a more readily accessible metric, the degree of task completion by a team, as a proxy. Social science literature (e.g., Kozlowski et al., 2006; Oliveira et al., 2014) suggest a correlation between group cohesion measures and task completion success, and we would like to see if our group cohesion model can correctly predict group task success. Moreover, if a low degree of group cohesion reduces the likelihood of task success, could a timely intervention "fix" the group dynamics to improve the chances of task

success? What's more, could such interventions be executed by strategically placed autonomous AI conversational agents? We are currently working towards finding answers to these and related questions.

ACKNOWLEDGEMENT

Research funding provided by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), through the Army Research Laboratory. Additional funding provided by the Defense Advanced Research Project Agency (DARPA), the Airforce Research Laboratory, and the Army Research Laboratory. All statements of fact, opinion or conclusions contained herein are those of the author and should not be construed as representing the official views or policies of IARPA, the ODNI, AFRL, ARL, DARPA, or the U.S. Government. This paper is based on the joint work of a number of people: Aaron Broadwell, Ting Liu, Samira Shaikh, Jennifer Stromer-Galley, Sarah Taylor, Nick Webb, Ning Sa, and many graduate students at UAlbany and RPI.

REFERENCES

Bales, R. F. (1950). A Set of Categories for the Analysis of Small Group Interaction. *American Sociological Review*, *15*(2), 257–263.

Bales, Robert Freed (2001) *Social Interaction Systems, Theory and Measurement*, Transaction Pub.

Broadwell, George Aaron, Umit Boz, Jennifer Stromer-Galley, Tomek Strzalkowski, Samira Shaikh, Sarah Taylor, Xiaoai Ren and Ting Liu. (2012) "Assessing Group Cohesion in Task-based Online Dialogues: A quantitative study." Georgetown University Round Table on Language & Linguistics, Washington, D. C.

George Aaron Broadwell, Jennifer Stromer-Galley, Tomek Strzalkowski, Sarah Taylor, Umit Boz, Samira Shaikh, Liu Ting, and Nick Webb (2013) Modeling Socio-Cultural Phenomena in Discourse. *Journal of Natural Language Engineering*. 19(2): 213–257. Cambridge.

Gregorios Katsios, Ning Sa, and Tomek Strzalkowski. "Social Convos: A New Approach to Modeling Information Diffusion in Social Media." Proceedings of 10th AHFE Conference, Washington, DC. Springer Nature. 2019.

Kozlowski, S. W. J., & Ilgen, D. R. "Enhancing the Effectiveness of Work Groups and Teams." *Psychological Science in the Public Interest*, *7*(3), 77–124. 2006.

Oliveira, A. W., Boz, U., Broadwell, G. A., & Sadler, T. D. (2014). Student leadership in small group science inquiry. *Research in Science & Technological Education*, *32*(3), 281–297.

S. Shaikh, T. Strzalkowski, A. Broadwell, J. Stromer-Galley, and N. Webb (2010) "MPC: A Multi-Party Chat Corpus For Modeling Social Phenomena in Discourse".

S. Shaikh, T. Strzalkowski, J. Stromer-Galley, G. Broadwell, T. Liu, and R. Mikeal Martey, "Modeling Leadership Behavior of Players in Virtual Worlds," *Proc. AAAI Conf. Artif. Intell. Interact. Digit. Entertain.*, vol. 11, no. 5, pp. 29–35, Jun. 2021, doi: 10.1609/aiide.v11i5.12844.

T. Strzalkowski et al., "Influence and Power in Group Interactions," in *Social Computing, Behavioral-Cultural Modeling and Prediction*, vol. 7812, A. M. Greenberg, W. G. Kennedy, and N. D. Bos, Eds., in Lecture Notes in Computer Science, vol. 7812., Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 19–27. doi: 10.1007/978–3-642–37210-0_3.

Tomek Strzalkowski, George Aaron Broadwell, Jennifer Stromer-Galley, Samira Shaikh, and Sarah Taylor "Modeling Leadership and Influence in Online Multi-Party Discourse." COLING-2012 Conference, Mumbai, India.